

# A Generalização dos Resultados de Estudos Clínicos Através da Análise de Subgrupos



## Generalization of the Results of Clinical Studies Through the Analysis of Subgroups

João COSTA, Filipa FARELEIRA, Raquel ASCENSÃO, António VAZ CARNEIRO

Acta Med Port 2012 Jan-Feb;25(1):44-50

### RESUMO

As análises de subgrupos em ensaios clínicos são habitualmente efectuadas para definir a potencial heterogeneidade do efeito do tratamento relacionado com o risco de base ou da fisiopatologia, com a aplicação prática da terapêutica ou ainda da subutilização da terapêutica na prática clínica de rotina devido a incerteza do benefício/risco desta. Quando apropriadamente planeadas, as análises de subgrupos são metodologias válidas para discernir benefícios terapêuticos em subgrupos de doentes, fornecendo desta maneira evidência de boa qualidade de suporte à decisão clínica. No entanto, para a sua realização ser correcta devem ser indicadas *a priori* no protocolo inicial, devem restringir-se ao *outcome* primário, devem ser feitos em número limitado, devem ser todas reportadas e, mais importante, devem ser submetidas a testes estatísticos de interacção.

Neste artigo apresentamos um exemplo no tratamento da osteoporose pós-menopáusica, em que o benefício de uma intervenção (quanto mais elevado o risco de fractura, maior o benefício do tratamento) com um agente específico (o bazedoxifeno) só foi esclarecido com uma análise *posthoc* de toda a amostra do estudo inicial.

### ABSTRACT

Subgroup analysis in clinical trials are usually performed to define the potential heterogeneity of treatment effect in relation with the baseline risk, physiopathology, practical application of therapy or the under-utilization in clinical practice of effective interventions due to uncertainties of its benefit/risk ratio. When appropriately planned, subgroup analysis are a valid methodology the define benefits in subgroups of patients, thus providing good quality evidence to support clinical decision making. However, in order to be correct, subgroup analysis should be defined *a priori*, done in small numbers, should be fully reported and, most important, must endure statistical tests for interaction.

In this paper we present an example of the treatment of post-menopausal osteoporosis, in which the benefits of an intervention (the higher the fracture risk is, the better the benefit is) with a specific agent (bazedoxifene) was only disclosed after a post-hoc analysis of the initial global trial sample.

### INTRODUÇÃO

A análise de subgrupos consiste na avaliação dos efeitos de uma dada intervenção terapêutica numa determinada variável de eficácia ou segurança em subgrupos de doentes definidos de acordo com características de base específicas.

As análises de subgrupos em grandes amostras de doentes de ensaios clínicos aleatorizados, prospectivos e controlados (*randomized controlled trials* – RCTs) apresentam uma atracção muito especial. Por exemplo, num RCT estudando os efeitos de uma dada estratégia terapêutica anti-hipertensiva, a identificação do benefício clínico dessa terapêutica nos doentes diabéticos em geral e em particular naqueles doentes diabéticos com ou sem doença coronária é obviamente de grande relevância clínica.<sup>1</sup> Da mesma forma, é clinicamente relevante saber se o benefício (redução do número de enfartes miocárdio e mortes) de uma estratégia invasiva, comparativamente a uma conservadora, em doentes com síndromas coronárias agudas sem elevação do segmento ST, é diferente de acordo com risco de base do doente.<sup>2</sup>

Frequentemente, a estimativa do potencial benefício em subgrupos de doentes clinicamente relevantes é apenas possível recorrendo a análises de subgrupos de grandes RCTs. Isto tem a ver com o facto da realização de mega-ensaios que especificamente incluam apenas este

tipo de doentes é (com poucas excepções) inviável e isto por várias razões, nomeadamente a dificuldade de recrutamento de um número significativo de sujeitos. Para além da estimativa do efeito de uma intervenção terapêutica em doentes com características específicas, a análise de subgrupos é também realizada para avaliar a consistência dos resultados em várias subpopulações.<sup>3,4</sup>

A validade externa dum RCT, isto é, a potencial generalização dos seus resultados a outros doentes não estudados no ensaio, baseia-se precisamente na capacidade dos resultados daqueles serem aplicáveis a pessoas com características individuais próprias, muitas vezes diferentes das da amostra estudada. Exemplos incluem a idade (idosos, crianças), género (mulheres), índice de massa corporal (obesos) ou co-morbilidades variadas.<sup>5</sup>

A análise de subgrupos, apesar de intuitiva e por vezes clinicamente relevante, levanta, no entanto, problemas específicos muito importantes<sup>3</sup> já que, se não for efectuada com regras muito rigorosas, poderá produzir resultados espúrios que nada têm a ver com a realidade.<sup>4,6,7</sup> Por exemplo, a aspirina reduz a mortalidade pós-enfarte de miocárdio comparativamente com o placebo (9,4% vs. 11,8%; *odds reduction*: 23 ± 4%;  $p < 0.00001$ ), mas uma análise de subgrupos em função dos signos do zodíaco clássico mostrou que nos doentes do signo gémeos e balança a aspirina tem um ligeiro risco acrescido de morte (9 ± 13%; não

Tabela 1 – Critérios de credibilidade de uma análise de subgrupos

DESENHO DO ESTUDO
A variável do subgrupo é uma característica que foi medida na linha de base, ou depois da aleatorização?
O efeito é sugerido por comparações efectuadas dentro da amostra do estudo ou entre estudos diferentes?
A hipótese foi especificada <i>a priori</i> ?
O sentido do efeito de subgrupo foi especificado <i>a priori</i> ?
O efeito de subgrupo foi um de entre um número limitado de efeitos hipotéticos testados?
ANÁLISE DO ESTUDO
O teste de interacção sugere uma baixa probabilidade que o efeito de subgrupo seja devido ao acaso?
O efeito de subgrupo é independente?
CONTEXTO
A dimensão do efeito de subgrupo é significativa?
A interacção é consistente de estudo para estudo?
A interacção é consistente entre resultados (outcomes) do estudo?
Existe evidência indirecta que suporte a interacção hipotética ( <i>rationale</i> biológico)?

significativo), enquanto que nos doentes com outros signos zodiacos a redução da mortalidade com aspirina era maior ( $28 \pm 5\%$ ;  $p < 0.00001$ )!<sup>8</sup>

O aspecto mais importante de uma análise de subgrupos é que ela, para ter validade externa, deve ser planeada antes do início do estudo, com critérios explícitos e incluída no protocolo inicial.<sup>3,4,9</sup> Ou seja, uma análise *post-hoc* de subgrupos (que é assim que se designa quando não foi pré-especificada, isto é, contemplada *ab initio* antes do estudo ter o seu início), não tem, à partida, a mesma validade, nomeadamente porque não se sabe os motivos porque foi realizada aquela e não outra, nem o número de outras análises *post-hoc* que não foram apresentadas por terem resultados negativos. Para além destas análises *post-hoc* de subgrupos, existem, no entanto, outras formas de análises *post-hoc* que, ao contrário destas análises de subgrupos, reavaliam posteriormente os resultados, mas considerando a totalidade dos doentes aleatorizados no RCT. A validade deste tipo de análise é obviamente maior, dado o significativo menor risco de enviesamento e de encontrar diferenças devido ao acaso.

Um outro aspecto metodológico importante nas análises de subgrupos é o número de análises realizadas, principalmente quando os resultados para a população global do estudo são negativos. À medida que são testados mais subgrupos de doentes, diminuimos o poder do estudo para mostrar diferenças e aumentamos a probabilidade de encontrar, apenas pelo acaso, resultados falsos positivos (quando a eficácia é idêntica, a probabilidade de encontrar um teste significativo por cada 10 explorados é de cerca de 40%).<sup>3,4</sup>

A análise de subgrupos deve ser realizada para o objectivo primário do estudo e recorrer a métodos estatísticos apropriados (nomeadamente a testes estatísticos de interacção) para avaliar se a dimensão do efeito da terapêutica é ou não diferente entre os subgrupos. Esta análise é diferente de realizar um teste estatístico para cada subgrupo.<sup>9</sup>

Neste artigo iremos apresentar e discutir a validade de vários tipos de análise de subgrupos, utilizando como exemplo o caso dos moduladores selectivos do receptor de estrogénio (raloxifeno e bazedoxifeno) no tratamento da osteoporose em mulheres pós-menopausa (OPM).

### Problemas metodológicos com análises de subgrupos

Os principais problemas metodológicos referentes à análise de subgrupos em RCTs foram anteriormente referidos. Não cabe num artigo deste tipo a discussão exaustiva destes problemas, sendo que o leitor interessado pode procurar informação detalhada em alguns textos clássicos.<sup>10-14</sup>

Recentemente foi publicado um conjunto de critérios para que uma análise de subgrupos seja credível.<sup>15</sup> Os autores actualizaram os sete critérios já publicados anteriormente,<sup>16</sup> adicionando quatro mais (Tabela 1). Para além desta tabela, existem outras fontes semelhantes, com recomendações análogas.<sup>17, 18</sup>

### Raloxifeno e Bazedoxifeno no Tratamento da OPM

A osteoporose é a doença óssea metabólica mais frequente e tem uma prevalência muito elevada. Trata-se de uma doença esquelética sistémica que se caracteriza pela diminuição da massa óssea e por uma alteração da qualidade microestrutural do osso, que levam a uma diminuição da resistência óssea e consequente aumento do risco de fracturas, sendo estas mais frequentes nas vértebras dorsais e lombares, extremidade distal do rádio e fémur proximal. Esta patologia pode ser considerada tanto uma doença silenciosa e sistémica, como um factor de risco progressivo para fractura. Na maioria das mulheres com OPM, a perda óssea relaciona-se principalmente com a reabsorção óssea excessiva.<sup>19,20</sup>

A definição operativa da OMS de osteoporose baseia-se essencialmente na quantificação da densidade mineral óssea (DMO) avaliada por absorciometria radiológica de dupla energia (DEXA).<sup>21</sup>

As opções terapêuticas incluem estratégias não farmacológicas - exercício físico moderado (marcha), cessação tabágica, redução do consumo de bebidas alcoólicas e uma dieta rica em cálcio e vitamina D – e farmacológicas – bifosfonatos, terapêutica hormonal de substituição, ranelato de estrôncio, teriparatida, PTH completa e calcitonina, os moduladores selectivos do receptor de estrogénio (raloxifeno e bazedoxifeno) e mais recentemente o denosumab.<sup>22</sup>

Uma distinção importante recai sobre o tipo de estratégia terapêutica: a prevenção ou o tratamento da osteoporose. No caso do tratamento da OPM, a eficácia do raloxifeno e o bazedoxifeno foram avaliados em 2 grandes RCTs de fase III metodologicamente semelhantes: o estudo MORE<sup>23</sup> e o estudo de Silverman *et al.*<sup>24</sup>

O estudo MORE<sup>23</sup> incluiu 7.705 doentes com OPM as quais foram estratificadas em 2 grupos pré-aleatorização: Grupo 1 - *T score* < -2,5; Grupo 2 - História de fractura vertebral com ou sem diminuição da DMO. Em cada grupo as doentes foram aleatorizadas para placebo, raloxifeno 60mg e 120mg. A variável de eficácia primária foi a incidência de fracturas vertebrais (por determinação radiológica) e a variação da DMO. Aos 36 meses, e para o total da população avaliada (Grupo 1 e Grupo 2), verificou-se para a população por intenção-de-tratar – ITT (n=6.828) uma redução de 30% e 50% do risco de fractura vertebral nos grupos do raloxifeno 60mg (6,6% vs. 10,1%; Risco Relativo: 0,7 [IC95%: 0,5-0,8]) e raloxifeno 120mg (5,4% vs. 10,1%; RR: 0,5 [IC95%: 0,4-0,7]), respectivamente, em comparação com o placebo. A dimensão do efeito foi semelhante nas doentes sem (Grupo 1; 66% das doentes) ou com (Grupo 2; 34% das doentes) história prévia de fractura vertebral. O raloxifeno, em comparação com o placebo, aumentou significativamente ( $p<0.001$  para todas as comparações) a DMO do colo do fémur entre 2,1% (60mg) e 2,4% (120mg) e na coluna vertebral entre 2,6% (60mg) e 2,7 (120mg). Não se verificou uma diferença significativa no risco de fractura não-vertebral no grupo do raloxifeno (60 e 120mg) em comparação com o placebo (RR: 0,9% [IC95%: 0,8-1,1]).

O estudo de Silverman *et al.*<sup>24</sup> incluiu 7.492 doentes aleatorizadas para bazedoxifeno (20mg e 40mg), raloxifeno (60mg) ou placebo. A variável de eficácia primária foi também a incidência de novas fracturas vertebrais ao fim de 36 meses. Verificou-se para a população ITT (n=6.847) uma redução do risco de fractura de 42% nos grupos bazedoxifeno 20mg (2,3% vs. 4,1%; *Hazard Ratio*: 0,58 [0,38-0,89]) e raloxifeno 60mg (2,3 vs. 4,1%; HR: 0,58 [0,38-0,89]), comparativamente ao placebo. Não houve diferenças significativas entre nenhum dos grupos de tratamento activo, quer nas doentes sem fracturas prévias (44% das doentes), quer nas doentes com fracturas prévias (56% das doentes). O bazedoxifeno e o raloxifeno, em comparação com o placebo, aumentaram significativamente ( $p<0.001$  para todas as comparações) a DMO do colo do fémur entre 0,27% (bazedoxifeno 20mg) e 0,90% (raloxifeno 60mg) e na coluna vertebral entre 2,21% (bazedoxifeno 20mg) e 2,96% (raloxifeno 60mg). A incidência de fracturas não vertebrais foi semelhante entre os grupos em estudo (bazedoxifeno 20

mg: 5,7%; bazedoxifeno 40 mg: 5,6%; raloxifeno 60 mg: 5,9%; placebo: 6,3%).

### Análise de Subgrupos

No estudo MORE<sup>23</sup> são apresentados os resultados de várias análises de subgrupos (idade, DMO, histerectomia prévia, terapêutica hormonal de substituição prévia) para a variável primária, não sendo claro que estas análises tenham sido pré-especificadas no protocolo (ou seja, que não sejam análises *post-hoc*), nem os métodos estatísticos utilizados.

No estudo de Silverman *et al.*<sup>24</sup> são apresentados os resultados de uma análise de subgrupos (existência ou não de fractura à entrada para o estudo) para a variável primária. Esta análise foi pré-especificada no protocolo do estudo e utilizou testes estatísticos de interacção ( $p=0,89$ ; ou seja, a dimensão do efeito do tratamento foi semelhante entre as doentes com ou sem fracturas prévias).

### Análise *Post-hoc* de Subgrupos

Sontag *et al.*<sup>25</sup> avaliaram recentemente o benefício do raloxifeno nas doentes com ou sem fracturas vertebrais prévias. Ao contrário da análise referida anteriormente para o bazedoxifeno,<sup>24</sup> esta análise não estava pré-especificada, tratando-se de uma análise *post hoc* de subgrupos dos resultados do ensaio MORE. Nesta análise, verificou-se que o efeito do raloxifeno na redução da incidência de fracturas vertebrais (variável de eficácia primária do estudo MORE) não diferiu significativamente entre as mulheres com (diferença absoluta vs. placebo: -8,2%;  $p<0.001$ ) ou sem (diferença absoluta vs. placebo: -2,8%;  $p<0.001$ ) fracturas vertebrais prévias (interacção não significativa;  $p=0,402$ ).

No estudo de Silverman *et al.*<sup>24</sup> são também apresentados os resultados de uma análise *post-hoc* de subgrupos que avaliou a eficácia do bazedoxifeno numa variável de eficácia secundária – risco de fracturas não vertebrais – num subgrupo de mulheres (n=1.772) com alto risco de fractura (sendo que o risco foi definido com base em factores de risco ósseos individuais conhecidos na *baseline*, nomeadamente: *T-score* colo do fémur  $\leq -3.0$  e/ou a presença de pelo menos uma fractura vertebral moderada ou grave ou múltiplas fracturas vertebrais ligeiras). Nesta análise, verificou-se para o bazedoxifeno 20 mg uma redução do risco de fracturas não vertebrais de 50% comparativamente com o placebo (HR: 0,50 [IC95%: 0,28-0,90];  $p=0.02$ ) e de 40% vs. o raloxifeno 60 mg (HR: 0,56 [IC95%: 0,31-1,01];  $p=0.05$ ).

Conforme mencionado previamente, a transposição para a prática clínica dos resultados deste tipo de análise deve ser realizada com cuidado face às limitações metodológicas e grau de incerteza dos resultados.

### Análise *Post-hoc* considerando toda a população aleatorizada

Conforme temos vindo a referir, as análises *post hoc* de subgrupos de doentes têm vários viéses, em particular a perda do poder estatístico. Neste contexto, uma possível

Tabela 2 - Risco de fractura osteoporótica major e do colo do fémur aos 10 anos na baseline

	Fractura major		Colo do fémur	
	Sem DMO	Com DMO	Sem DMO	Com DMO
<b>RALOXIFENO (n)</b>	7.703	7.652	7.703	7.652
<b>Média±DP</b>	18,7±10,8	20,6±11,0	5,7±5,4	6,6±6,2
<b>BAZEDOXIFENO (n)</b>	5.632	5.213	5.632	5.213
<b>Média±DP</b>	11,0±8,4	10,5±8,2	3,5±4,5	3,0±4,6

DP: Desvio-Padrão

Fonte: adaptado de Kanis et al. 2009<sup>26</sup> e 2010<sup>27</sup>.

abordagem é a análise, não de apenas um subgrupo de doentes, mas de toda a população aleatorizada.

Recentemente foram publicados os resultados de duas reanálises dos dados dos estudos MORE<sup>26</sup> e Silverman et al.<sup>27</sup> com base em modelos multivariados de regressão logística e que incluíram a totalidade da população aleatorizada nestes estudos. O objectivo foi o de avaliar a eficácia do raloxifeno e do bazedoxifeno na prevenção de todas as fracturas clínicas e das fracturas vertebrais morfométricas em função do risco fracturário de base das doentes estimado pelo algoritmo FRAX® (Fracture Risk Assessment tool). O algoritmo FRAX® (com dados da DMO) permite estimar a probabilidade de fractura osteoporótica major<sup>28</sup> – métrica clinicamente relevante de definição de alto risco – como variável contínua, tal como exigido pelo CHMP (Committee for Medicinal Products for Human Use) para os novos ensaios de fase III.<sup>29</sup> A hipótese avaliada foi de que a redução do risco de fractura com o raloxifeno e o bazedoxifeno, comparativamente ao placebo, era tanto maior quanto maior a probabilidade de fractura prévia à entrada no estudo e de que este efeito seria semelhante nas fracturas clínicas.

Quer em relação ao raloxifeno, quer em relação ao bazedoxifeno, e para maximizar o poder estatístico, os grupos do raloxifeno (60 e 120mg) no estudo MORE e os grupos do bazedoxifeno (20 e 40mg) no estudo Silverman et al. foram combinados, uma vez que não se verificaram diferenças significativas para qualquer *outcome* entre estes grupos de tratamento.

Quanto aos resultados para o estudo MORE, o total de doentes avaliadas nesta reanálise foi de 7.705, tendo sido possível estimar a probabilidade de fractura osteoporótica major num período de 10 anos em 7.703 doentes (sem incorporar o valor da DMO) e em 7.652 doentes (incorporando o valor da DMO na estimativa da probabilidade de fractura) (Tabela 2).

Do total de doentes, 825 tiveram fracturas clínicas (677 não-vertebrais e 148 vertebrais). O raloxifeno foi associado a uma redução significativa de 42% na incidência de fracturas vertebrais morfométricas (HR: 0,58 [IC95%: 0,48-0,69];  $p<0.001$ ) e de 18% no total de fracturas clínicas (HR: 0,82 [IC95%: 0,71-0,95];  $p=0.006$ ), quando comparado com o placebo. Relativamente ao efeito sobre as fracturas clínicas, quando se excluíram as fracturas vertebrais, verificou-se uma redução não significativa de 9% (HR: 0,91; [IC95%: 0,78-1,07];  $p=0.26$ ).

O *Hazard Ratio* entre os grupos de tratamento (raloxifeno vs. placebo) não variou de forma significativa de acordo com a probabilidade de fractura, quer no caso das fracturas vertebrais morfométricas, quer no caso das fracturas clínicas. Na verdade, a redução do risco, quer de fracturas vertebrais morfométricas, quer de fracturas clínicas, foi menor nos grupos de maior risco fracturário (Tabela 3 e Tabela 4). No caso das fracturas vertebrais morfométricas, verificou-se que a eficácia diminuía de forma significativa com o aumento da idade ( $p=0.020$ ).

Quanto aos resultados para o estudo Silverman et al., o total de doentes avaliadas nesta reanálise foi de 5.643, tendo sido possível estimar a probabilidade de fractura osteoporótica major num período de 10 anos em 5.213 doentes (incorporando o valor da DMO na estimativa da probabilidade de fractura) (Tabela 2).

Durante o estudo, 412 doentes tiveram fracturas clínicas (39 vertebrais e 373 não vertebrais) e verificaram-se 164 novas fracturas vertebrais morfométricas. O bazedoxifeno foi associado a uma redução significativa (39%) na incidência de fracturas vertebrais morfométricas (HR: 0,61 [IC 95%: 0,43-0,86];  $p=0.005$ ) e a uma redução não significativa de 16% no total de fracturas clínicas (HR: 0,84 [IC 95%: 0,67-1,06];  $p=0.14$ ), quando comparado com o placebo.

O *Hazard Ratio* entre os grupos de tratamento (bazedoxifeno vs. placebo) variou de acordo com a probabilidade de fractura, quer no caso das fracturas vertebrais morfométricas, quer no caso das fracturas clínicas (Tabela 3 e Tabela 4).

Os *Hazard Ratios* foram significativos a partir do percentil 25 no caso das fracturas vertebrais morfométricas e a partir do percentil 75 no caso das fracturas clínicas. O valor de probabilidade de fractura a partir do qual se verificou um efeito estatisticamente significativo com bazedoxifeno foi de 6,9% (equivalente ao percentil 41) e 16% (equivalente ao percentil 80), no caso das fracturas vertebrais morfométricas e fracturas clínicas, respectivamente.

Os autores desta análise *post-hoc* do estudo Silverman et al. realizaram uma análise de sensibilidade considerando a probabilidade de fractura do colo do fémur em vez da probabilidade de fractura osteoporótica major (colo do fémur, fractura clínica vertebral, do antebraço ou do úmero). Os resultados foram semelhantes aos da análise primária considerando a probabilidade de fractura osteoporótica

Tabela 3 - Hazard Ratio e probabilidade a 10 anos de fracturas vertebrais morfométricas.

Fracturas vertebrais morfométricas				
Raloxifeno			Bazedoxifeno	
Percentil	Probabilidade (%)	HR (IC95%)	Probabilidade (%)	HR (IC95%)
10	8,4	0,48 (0,35-0,64)	2,8	0,73 (0,45-1,18)
25	13,3	0,50 (0,39-0,64)	4,5	0,71 (0,45-1,10)
50	21,1	0,54 (0,44-0,65)	8,2	0,65 (0,45-0,95)
75	30,3	0,59 (0,49-0,70)	14,0	0,58 (0,41-0,82)
90	40,1	0,64 (0,51-0,81)	21,7	0,49 (0,31-0,79)

Fonte: adaptado de Kanis et al. 2009<sup>26</sup> e 2010<sup>27</sup>.

Tabela 4 - Hazard Ratio e probabilidade a 10 anos de fracturas clínicas.

Fracturas clínicas				
Raloxifeno			Bazedoxifeno	
Percentil	Probabilidade (%)	HR (IC95%)	Probabilidade (%)	HR (IC95%)
10	8,4	0,74 (0,59-0,92)	2,8	1,02 (0,74-1,40)
25	13,3	0,76 (0,63-0,91)	4,5	0,98 (0,73-1,32)
50	21,1	0,80 (0,69-0,92)	8,3	0,91 (0,71-1,17)
75	30,3	0,84 (0,72-0,99)	14,5	0,80 (0,63-1,02)
90	40,1	0,90 (0,72-1,12)	22,4	0,68 (0,49-0,93)

Fonte: adaptado de Kanis et al. 2009<sup>26</sup> e 2010<sup>27</sup>.

Tabela 5 - Hazard Ratio e probabilidade de fractura do colo do fémur

Fracturas vertebrais morfométricas				Fracturas clínicas		
Percentil	Probabilidade (%)	HR	IC 95%	Probabilidade (%)	HR	IC 95%
10	0,2	0,69	0,46-1,03	0,2	0,95	0,73-1,24
25	0,5	0,68	0,46-1,00	0,5	0,94	0,72-1,22
50	1,5	0,65	0,45-0,94	1,5	0,90	0,70-1,16
75	3,6	0,60	0,43-0,85	3,6	0,83	0,66-1,05
90	7,6	0,52	0,34-0,80	7,6	0,71	0,53-0,95

Fonte: adaptado de Kanis et al. 2009<sup>27</sup>.

major. A tabela 5 mostra os resultados desta análise.

De acordo com os resultados desta reanálise do estudo Silverman et al. (2008), os autores concluíram que, pese embora o bazedoxifeno não tenha um efeito global significativo na redução das fracturas clínicas, verifica-se uma redução significativa em doentes com alta probabilidade de fractura osteoporótica major. Em doentes com uma probabilidade  $\geq 16\%$  (correspondente ao percentil 80 dos doen-

tes avaliados no estudo), o bazedoxifeno está associado a uma redução significativa de 23% do risco de fracturas clínicas.

## DISCUSSÃO

As análises de subgrupos em RCTs são habitualmente efectuadas por diversas razões, das quais se destacam as seguintes:<sup>18</sup> 1) potencial heterogeneidade do efeito do



tratamento relacionado com o risco de base; 2) potencial heterogeneidade do efeito do tratamento relacionado com a fisiopatologia; 3) questões clinicamente importantes relacionadas com a aplicação prática da terapêutica; 4) subutilização da terapêutica na prática clínica de rotina devido a incerteza do benefício/risco da terapêutica.

Estas análises podem classificar-se de pré-especificadas (quando são planeadas e documentadas antes da análise final dos resultados, idealmente incluídas no protocolo) e *post-hoc* (em que as hipóteses testadas não são especificadas antes da análise dos dados).

Quando apropriadamente planeadas, as análises de subgrupos são metodologias válidas para discernir benefícios terapêuticos em subgrupos de doentes, fornecendo desta maneira evidência de boa qualidade de suporte à decisão clínica. Por vezes fazem sentido sob o ponto de vista biológico, ou são impostos pelos *sponsors* do estudo, pelos profissionais ou, por vezes, pelo próprio público. Se se realizarem, devem ser indicados no protocolo inicial, devem restringir-se ao *outcome* primário, devem ser feitos em número limitado, devem ser todos reportados e, mais importante, devem ser submetidos a testes estatísticos de interacção.<sup>3,4,9</sup>

Recorrendo ao exemplo dos moduladores selectivos do receptor de estrogénio (raloxifeno e bazedoxifeno) no tratamento da osteoporose em mulheres pós-menopausa apresentámos os vários tipos de análise de subgrupos (pré-especificadas e *posthoc*), assim como de análises *posthoc* da totalidade das amostras dos estudos em que foram analisados os impactos destas terapêuticas anti-osteoporóticas, através de estratificação de risco utilizando uma das mais bem validadas escalas para esta patologia (FRAX®), e que permitiu sugerir existir uma diferença no benefício do bazedoxifeno – não óbvia nas análises iniciais – em que este fármaco se relaciona com o risco de base de fractura: quanto mais elevado o risco de fractura, maior o benefício

do tratamento com bazedoxifeno – sendo as mulheres com OPM e elevado risco de fractura potencialmente o alvo terapêutico preferencial desta intervenção (devido ao facto do fármaco reduzir significativamente o risco de fracturas clínicas não vertebrais).

## CONCLUSÕES

A tensão existente entre a prática clínica – em que o médico necessita de identificar doentes com características específicas, próprias (os chamados subgrupos de risco) – e a investigação clínica – em que são estudados grandes grupos de doente homogêneos, com características semelhantes – foi já designada por um conhecido autor como uma “tragédia clínico-estatística”.<sup>30</sup> Esta afirmação baseou-se no conceito (estatístico) que uma análise *posthoc* é frequentemente inválida pelas razões já apresentadas, e no conceito (clínico) de que os doentes possuem características próprias que determinam a sua abordagem específica.

A ultrapassagem desta dicotomia passa por uma qualquer análise de subgrupos, mesmo quando realizada *posthoc*, poder em casos seleccionados ser de grande utilidade clínica e validade estatística, se realizada com rigor. Neste artigo apresentámos precisamente um destes casos, em que o benefício de uma intervenção (quanto mais elevado o risco de fractura, maior o benefício do tratamento) com um agente específico (o bazedoxifeno) só foi esclarecido com uma análise *posthoc* de toda a amostra do estudo inicial. Obviamente que estes resultados necessitam ser confirmados em futuros RCTs que incluam especificamente doentes de elevado risco de fractura estimado por FRAX®.

## CONFLITO DE INTERESSES

Não declarado.

## FONTES DE FINANCIAMENTO

Não declaradas.

## REFERÊNCIAS

- Cooper-DeHoff RM, Gong Y, Handberg EM, Bavry AA, Denardo SJ, Bakris GL et al. Tight blood pressure control and cardiovascular outcomes among hypertensive patients with diabetes and coronary artery disease. JAMA 2010;304(1):61-68.
- Fox KA, Poole-Wilson P, Clayton TC, Henderson RA, Shaw TR, Wheatley DJ et al. 5-year outcome of an interventional strategy in non-ST-elevation acute coronary syndrome: the British Heart Foundation RITA 3 randomised trial. Lancet 2005;366(9489):914-920.
- Schulz KF, Grimes DA. Multiplicity in randomised trials II: subgroup and interim analyses. Lancet 2005;365(9471):1657-1661.
- Lagakos SW. The challenge of subgroup analyses-reporting without distorting. N Engl J Med 2006;354(16):1667-1669.
- Carneiro AV. Análise de sub-grupos em ensaios clínicos terapêuticos. Rev Port Cardiol 2002;21:339-346.
- Bhatt DL, Fox KA, Hacke W, Berger PB, Black HR, Boden WE et al. Clopidogrel and aspirin versus aspirin alone for the prevention of atherothrombotic events. N Engl J Med 2006;354(16):1706-1717.
- Pfeffer MA, Jarcho JA. The charisma of subgroups and the subgroups of CHARISMA. N Engl J Med 2006;354(16):1744-1746.
- Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Lancet 1988;2(8607):349-360.
- Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine--reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357(21):2189-2194.
- User's Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice. 1st ed. Chicago: JAMA & Archives Journals AMA; 2001.
- Mayer D. Essential evidence-based medicine. 1st ed. Cambridge: Cambridge University Press 2004.
- Advances in clinical trial biostatistics. 1st ed. New York: Marcel Dekker, Inc.; 2004.
- Elwood M. Critical appraisal of epidemiological studies and clinical trials. 3d ed. Oxford: Oxford University Press 2007.
- Haynes RB, Sackett DL, Guyatt GH, Tugwell P. Clinical epidemiology. 3d ed. Philadelphia: Lippincott Williams & Wilkins 2006.
- Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ 2010;340:850-854.
- Oxman AD, Guyatt GH. A consumer's guide to subgroup analysis. Ann Intern Med 1992;116(1):78-84.
- Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine--reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357(21):2189-2194.
- Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 2005;365(9454):176-186.

19. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. *Lancet* 2002;359(9321):1929-1936.
20. Cummings SR, Melton L.J. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761-1767.
21. <http://www.dgs.pt/upload/membro.id/ficheiros/i006345.pdf> [Acedido em 26 de Setembro de 2011].
22. Kanis JA, Burlet N, Cooper C, Delmas PD, Reginster JY, Borgstrom F *et al.* European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos Int* 2008;19(4):399-428.
23. Ettinger B, Black DM, Mitlak BH, Knickerbocker RK, Nickelsen T, Genant HK *et al.* Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. *JAMA* 1999;282(7):637-645.
24. Silverman SL, Christiansen C, Genant HK, Vukicevic S, Zanchetta JR, de Villiers TJ *et al.* Efficacy of bazedoxifene in reducing new vertebral fracture risk in postmenopausal women with osteoporosis: results from a 3-year, randomized, placebo-, and active-controlled clinical trial. *J Bone Miner Res* 2008;23(12):1923-1934.
25. Sontag A, Wan X, Krege JH. Benefits and risks of raloxifene by vertebral fracture status. *Curr Med Res Opin* 2010;26(1):71-76.
26. Kanis JA, Johansson H, Oden A, McCloskey EV. A meta-analysis of the efficacy of raloxifene on all clinical and vertebral fractures and its dependency on FRAX. *Bone* 2010;47(4):729-735.
27. Kanis JA, Johansson H, Oden A, McCloskey EV. Bazedoxifene reduces vertebral and clinical fractures in postmenopausal women at high risk assessed with FRAX. *Bone* 2009;44(6):1049-1054.
28. Kanis JA, Oden A, Johansson H, Borgström F, Ström O, McCloskey E. FRAX and its applications to clinical practice. *Bone* 2009;44(5):734-743.
29. Committee for Medicinal Products for Human Use (CHMP). Guideline on the evaluation of medicinal products in the treatment of primary osteoporosis. Ref CPMP/EWP/552/95Rev.2. London: CHMP; 2006 [Nov].
30. Feinstein AR. The problems of cogent subgroups: a clinicostatistical tragedy. *J Clin Epidemiol* 1998;51(4):297-299.